

RE P U B L I Q U E F R A N C A
Rec'd PCT/PTO 28 JUN 2004



PCT/FR 02/04549

10/500441

REC'D 18 MAR 2003

WIPO

PCT

BREVET D'INVENTION

CERTIFICAT D'UTILITÉ - CERTIFICAT D'ADDITION

COPIE OFFICIELLE

Le Directeur général de l'Institut national de la propriété industrielle certifie que le document ci-annexé est la copie certifiée conforme d'une demande de titre de propriété industrielle déposée à l'Institut.

Fait à Paris, le 31 JAN. 2003

Pour le Directeur général de l'Institut
national de la propriété industrielle
Le Chef du Département des brevets

Martine PLANCHE

DOCUMENT DE PRIORITÉ

PRÉSENTÉ OU TRANSMIS
CONFORMÉMENT À LA
RÈGLE 17.1.a) OU b)

INSTITUT
NATIONAL DE
LA PROPRIÉTÉ
INDUSTRIELLE

SIEGE
26 bis, rue de Saint Petersburg
75800 PARIS cedex 08
Téléphone : 33 (0)1 53 04 53 04
Télécopie : 33 (0)1 53 04 45 23
www.inpi.fr

BEST AVAILABLE COPY



26 bis, rue de Saint Pétersbourg
75800 Paris Cedex 08
Téléphone : 01 53 04 53 04 Télécopie : 01 42 94 86 54

BREVET D'INVENTION

CERTIFICAT D'UTILITÉ

Code de la propriété intellectuelle - Livre VI



N° 11354*01

REQUÊTE EN DÉLIVRANCE 1/2

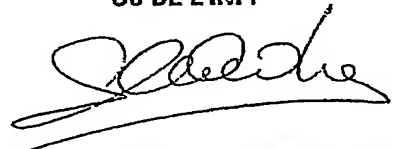
Important Remplir impérativement la 2ème page.

Cet imprimé est à remplir lisiblement à l'encre noire

DB 540 W / 190500

REMISE DES PIÈCES DATE 27 DEC 2001 LIEU 75 INPI PARIS N° D'ENREGISTREMENT 0116949 NATIONAL ATTRIBUÉ PAR L'INPI DATE DE DÉPÔT ATTRIBUÉE 27 DEC. 2001 PAR L'INPI		1 NOM ET ADRESSE DU DEMANDEUR OU DU MANDATAIRE À QUI LA CORRESPONDANCE DOIT ÊTRE ADRESSÉE Cabinet BALLOT 7, rue Le Sueur 75116 PARIS	
Vos références pour ce dossier (facultatif) 015784 VG/CC			
Confirmation d'un dépôt par télécopie <input type="checkbox"/> N° attribué par l'INPI à la télécopie			
2 NATURE DE LA DEMANDE		Cochez l'une des 4 cases suivantes	
Demande de brevet		<input checked="" type="checkbox"/>	
Demande de certificat d'utilité		<input type="checkbox"/>	
Demande divisionnaire		<input type="checkbox"/>	
Demande de brevet initiale		N°	Date
ou demande de certificat d'utilité initiale		N°	Date
Transformation d'une demande de brevet européen		N°	Date
Demande de brevet initiale		N°	Date
3 TITRE DE L'INVENTION (200 caractères ou espaces maximum) PROCÉDE DE CARACTÉRISATION D'UN SIGNAL SONORE			
4 DÉCLARATION DE PRIORITÉ OU REQUÊTE DU BÉNÉFICE DE LA DATE DE DÉPÔT D'UNE DEMANDE ANTÉRIEURE FRANÇAISE		Pays ou organisation Date N° Pays ou organisation Date N° Pays ou organisation Date N° <input type="checkbox"/> S'il y a d'autres priorités, cochez la case et utilisez l'imprimé «Suite»	
5 DEMANDEUR		<input type="checkbox"/> S'il y a d'autres demandeurs, cochez la case et utilisez l'imprimé «Suite»	
Nom ou dénomination sociale		FRANCE TELECOM	
Prénoms			
Forme juridique		S.A. (Société Anonyme)	
N° SIREN			
Code APE-NAF			
Adresse	Rue	6, Place d'Alleray	
	Code postal et ville	75015 PARIS	
Pays		FRANCE	
Nationalité			
N° de téléphone (facultatif)			
N° de télécopie (facultatif)			
Adresse électronique (facultatif)			

1/1

REMISE DES PIÈCES DATE 27 DEC 2001 LIEU 75 INPI PARIS N° D'ENREGISTREMENT 0116940 NATIONAL ATTRIBUÉ PAR L'INPI		Réservé à l'INPI		DB 540 W / 100500	
Vos références pour ce dossier : (facultatif)			015784 VG/CC		
6 MANDATAIRE					
Nom			BALLOT		
Prénom			Paul		
Cabinet ou Société			Cabinet BALLOT		
N° de pouvoir permanent et/ou de lien contractuel					
Adresse	Rue		7, rue Le Sueur		
	Code postal et ville		75116 PARIS		
N° de téléphone (facultatif)			01.40.67.11.99		
N° de télécopie (facultatif)			01.45.01.98.28		
Adresse électronique (facultatif)					
7 INVENTEUR (S)					
Les inventeurs sont les demandeurs			<input type="checkbox"/> Oui <input checked="" type="checkbox"/> Non Dans ce cas fournir une désignation d'inventeur(s) séparée		
8 RAPPORT DE RECHERCHE			Uniquement pour une demande de brevet (y compris division et transformation)		
Établissement immédiat ou établissement différé			<input checked="" type="checkbox"/> <input type="checkbox"/>		
Paiement échelonné de la redevance			Paiement en deux versements, uniquement pour les personnes physiques <input type="checkbox"/> Oui <input type="checkbox"/> Non		
9 RÉDUCTION DU TAUX DES REDEVANCES			Uniquement pour les personnes physiques <input type="checkbox"/> Requête pour la première fois pour cette invention (joindre un avis de non-imposition) <input type="checkbox"/> Requête antérieurement à ce dépôt (joindre une copie de la décision d'admission pour cette invention ou indiquer sa référence) :		
Si vous avez utilisé l'imprimé «Suite», indiquez le nombre de pages jointes					
10 SIGNATURE DU DEMANDEUR OU DU MANDATAIRE (Nom et qualité du signataire) BALLOT Paul N° 92-1009 Cabinet BALLOT			VISA DE LA PRÉFECTURE OU DE L'INPI 		

PROCEDE DE CARACTERISATION D'UN SIGNAL SONORE

L'invention concerne un procédé de caractérisation selon des paramètres spécifiques, d'un signal sonore évoluant selon le temps dans différentes bandes de fréquences.

5 Le domaine de l'invention est celui de la reconnaissance de signaux sonores appliquée en particulier à l'identification d'œuvres musicales utilisées sans autorisation.

10 En effet, le développement des techniques de numérisation et du multimédia a provoqué une augmentation considérable de telles utilisations frauduleuses. Il en résulte une nouvelle difficulté pour les organismes chargés de collecter les droits d'auteur, puisqu'il faut être en mesure d'identifier
15 ces utilisations en particulier sur les réseaux numériques interactifs tel Internet, afin d'effectuer de façon satisfaisante la perception et la répartition des rémunérations dues aux auteurs de ces œuvres musicales.

20 Dans la suite, pour ne pas se limiter aux œuvres musicales, on considérera plus généralement un signal sonore.

25 Le but de la présente invention est donc de constituer une base de données de signaux sonores, chaque signal sonore étant caractérisé par une empreinte de sorte que, étant donné un signal sonore inconnu que l'on caractérise de la même façon, on puisse effectuer une recherche et une comparaison

rapides de l'empreinte de ce signal inconnu avec l'ensemble des empreintes de la base de données.

L'empreinte est constituée de paramètres spécifiques déterminés de la façon suivante.

5 On décompose dans un premier temps le signal sonore dont l'amplitude $x(t)$ varie avec le temps t , selon différentes bandes de fréquences k : $x(k,t)$ est l'amplitude du signal sonore filtré dans la bande de fréquence k et représenté figure 1a).

10 On calcule, comme représenté figure 1c) l'énergie à court terme $E(k,t)$ de ce signal sonore filtré à l'aide d'une fenêtre $h(t)$ représentée figure 1b), ayant un support de $2N$ secondes. On réitère ce calcul en faisant glisser cette fenêtre toutes les S secondes.

15 Ces valeurs $E(k,t)$ constituent des paramètres spécifiques d'un extrait de $2N$ secondes du signal sonore $x(k,t)$ dans la bande de fréquences k .

On peut obtenir d'autres paramètres en calculant pour différentes bandes de fréquences j , l'énergie de $E(k,t)$ à l'aide d'une fenêtre $h'(t)$ représentée figure 20 2b), ayant un support de $2N'$ secondes ; on réitère ce calcul en faisant glisser cette fenêtre toutes les S' secondes : on obtient $F(j,k,t)$ représenté figure 2c). On normalise les valeurs $F(j,k,t)$ par rapport à leur 25 maximum de façon à les rendre indépendantes de l'amplitude du signal sonore.

Ainsi normalisées, ces valeurs constituent des paramètres spécifiques d'un extrait de $2N'$ secondes du signal sonore $x(k,t)$ dans la bande de fréquences k .

30 On peut également calculer la phase de $E(k,t)$ pour différentes bandes de fréquences j : on obtient

$P(j,k,t)$. On normalise les valeurs $P(j,k,t)$ par rapport à une valeur de référence $P(1,j,t)$ et on obtient alors d'autres paramètres spécifiques d'un extrait de $2N'$ secondes du signal sonore.

- 5 D'autres paramètres peuvent être ajoutés tels que la valeur moyenne de l'énergie $E(k,t)$.

L'invention a pour objet un procédé de caractérisation selon des paramètres spécifiques, d'un
 10 signal sonore $x(t)$ évoluant selon le temps t pendant une durée D dans différentes bandes de fréquences k et alors noté $x(k,t)$, principalement caractérisé en ce qu'il consiste à mémoriser le signal $x(t)$, à calculer l'énergie $E(k,t)$ dudit signal $x(k,t)$ pour chacune
 15 desdites bandes de fréquences k , k variant de 1 à K et selon une fenêtre temporelle $h(t)$ d'une durée $2N$, mémoriser les valeurs de l'énergie $E(k,t)$ obtenues, ces valeurs constituant les paramètres spécifiques d'un extrait d'une durée $2N$ du signal sonore $x(t)$ et à
 20 réitérer ce calcul à intervalles réguliers pour obtenir l'ensemble des paramètres spécifiques pour la durée D du signal sonore $x(t)$.

Il consiste en outre à calculer et mémoriser l'énergie $F(k,j,t)$ de $E(k,t)$ pour des bandes de
 25 fréquences j , j variant de 1 à J , selon une fenêtre temporelle $h'(t)$ d'une durée $2N'$, les $J \times K$ valeurs de l'énergie $F(j,k,t)$ obtenues constituant les paramètres spécifiques d'un extrait d'une durée $2N'$ du signal sonore $x(t)$ et à réitérer ce calcul à intervalles

réguliers pour obtenir l'ensemble des paramètres spécifiques pour la durée D du signal sonore $x(t)$.

Il consiste éventuellement à calculer la phase $P(j,k,t)$ de l'énergie $E(k,t)$ pour des bandes de fréquences j , j variant de 1 à J , j différent de k , et à inclure les valeurs de la phase $P(j,k,t)$ obtenues parmi les paramètres spécifiques du signal sonore $x(t)$.

Il peut aussi consister à calculer pour chaque bande de fréquence j , la valeur moyenne de l'énergie $E(k,t)$ sur $2N'$ secondes, à réitérer ce calcul à intervalles réguliers pour obtenir l'ensemble des paramètres spécifiques pour la durée D du signal sonore $x(t)$ et à inclure les valeurs moyennes obtenues parmi les paramètres spécifiques du signal sonore $x(t)$.

Selon une caractéristique, il consiste à considérer les paramètres spécifiques d'un signal sonore $x(t)$ comme les composantes d'un vecteur représentatif de $x(t)$, à positionner les vecteurs dans un espace à autant de dimensions que de paramètres, à définir des classes regroupant les vecteurs les plus proches et à enregistrer lesdites classes.

Les classes présentant des distances inter-classes et des distances intra-classes, le procédé consiste avantageusement à sélectionner parmi les paramètres spécifiques, les paramètres permettant d'obtenir des distances inter-classes relativement grandes devant les distances intra-classes et à enregistrer les paramètres sélectionnés.

L'invention concerne également un dispositif d'identification d'un signal sonore, caractérisé en ce qu'il comprend un serveur de base de données comprenant

des moyens pour la mise en œuvre du procédé de caractérisation d'un signal sonore selon des paramètres spécifiques, tel que décrit précédemment et des moyens de recherche dudit signal sonore dans la base de données.

De préférence, les moyens de recherche comprennent des moyens de reconnaissance directe de la classe à laquelle ledit signal sonore appartient et des moyens de recherche de la classe par comparaison des paramètres spécifiques du signal sonore inconnu avec ceux de la base de donnée, la classe étant choisie, par exemple, par la méthode des plus proches voisins.

D'autres particularités et avantages de l'invention apparaîtront clairement à la lecture de la description faite à titre d'exemple non limitatif et en regard des dessins annexés sur lesquels :

les figures 1a), 1b) et 1c) représentent respectivement des courbes schématiques de variation d'un signal sonore $x(k_i, t)$ filtré dans une bande de fréquences k_i , d'une fenêtre de Hamming $h(t)$ et de l'énergie $E(k_i, t)$ à court terme du signal $x(k_i, t)$,

les figures 2a), 2b) et 2c) représentent respectivement des courbes schématiques de variation de l'énergie $E(k_i, t)$ pour la bande de fréquence k_i , d'une fenêtre de Hamming $h'(t)$ et de l'énergie $F(j_m, k_i, t)$ de $E(k_i, t)$ pour la bande de fréquences j_m ,

la figure 3 illustre schématiquement un ensemble de vecteurs $V[x(t)]$ constituant l'empreinte d'un signal $x(k, t)$,

la figure 4 illustre schématiquement le stockage d'empreintes,

la figure 5 représente une classification des signaux sonores selon deux paramètres,

5 la figure 6 illustre une méthode de recherche d'un signal sonore par la méthode du plus proche voisin.

la figure 7 représente schématiquement un serveur de base de données de stockage des empreintes de signaux sonores.

10

Les signaux sonores que l'on traite selon ce procédé de caractérisation sont des signaux sonores enregistrés, notamment sur des disques compacts (Compact Disk en anglais).

15 On va considérer par la suite que le signal sonore $x(t)$ est un signal numérique échantillonné à une fréquence d'échantillonnage f_e , par exemple 11025 Hz correspondant au quart de la fréquence d'échantillonnage courante des disques compacts qui est
20 de 44100 Hz.

On peut cependant caractériser un signal sonore analogique : il faut au préalable le convertir en un signal numérique au moyen d'un convertisseur analogique-numérique.

25 Le signal sonore $x(k,t)$ représenté figure 1a) pour $k=k_i$ est donc un signal numérique échantillonné à la fréquence f_e et obtenu après un filtrage dans une bande de fréquences k_i . Chaque valeur de ce signal numérique échantillonné est codée par exemple sur 16 bits. Les
30 bandes de fréquences sont des bandes du spectre audible variant d'environ 20 Hz à 20kHz et découpé en K (k

varie de 1 à K) bandes de fréquences, K=127 par exemple.

L'énergie à court terme $E(k,t)$ représentée figure 1c) pour $k=k_i$, est calculée sur une fenêtre $h(t)$ de 2N secondes, par exemple une fenêtre de Hamming ayant un support d'environ 23 ms représentée figure 1b).

$E(k,t)$ est le carré du module d'une transformée du signal sonore échantillonné $x(t)$ dans le plan temps-fréquence ou dans le plan temps-échelle. Parmi les transformées utilisables, figurent la transformée de Fourier, la transformée en cosinus, la transformée de Hartley et la transformée en ondelettes. Un banc de filtres passe-bandes réalise également une telle transformée. La transformée de Fourier à court terme permet une représentation temps-fréquence adaptée à l'analyse des signaux musicaux. Ainsi l'énergie $E(k,t)$ s'écrit :

$$E(k,t) = \left| \sum_{n=-N}^{n=N} x(t + n/f_0) \cdot h(n/f_0) \cdot e^{-4i\pi kn/N} \right|^2$$

avec i tel que $i^2 = -1$

On fait glisser la fenêtre sur le signal sonore toutes les S secondes, par exemple toutes les 10 ms. $E(k,t)$ sera ainsi échantillonné toutes les 10 ms : on aura $E(k,t_0)$ puis $E(k,t_1)$ avec $t_1=t_0+10$ ms etc.

On code ainsi toutes les S secondes le signal sonore $x(t)$ par un vecteur à K composantes $E(k,t)$, chacune de ces composantes codant l'énergie de 23 ms du signal sonore $x(t)$ dans K bandes de fréquences.

On obtient d'autres paramètres en reproduisant en quelque sorte les calculs précédents et en les appliquant cette fois à $E(k,t)$ comme représenté figure 2a) à figure 2c).

5 On filtre l'énergie $E(k,t)$ dans J différentes bandes de fréquence : $E(j,k,t)$ est l'énergie $E(k,t)$ filtrée dans la bande de fréquences j , j variant de 1 à J avec par exemple $J=51$.

10 On calcule alors $F(j,k,t)$ représentée figure 2c) pour $k=k_1$ et $j=j_m$, sur une fenêtre $h'(t)$ de $2N'$ secondes, par exemple une fenêtre de Hamming ayant un support de 10 s. Ainsi, on peut écrire, avec i tel que $i^2 = -1$:

$$F(j,k,t) = \left| \sum_{n=-N'}^{n=N'} E(k,t+n/f_c) \cdot h'(n/f_c) \cdot e^{-4i\pi j n/N'} \right|^2$$

15

Dans notre exemple, toutes les secondes ($S'=1$), le signal sonore $x(t)$ est codé par 127×51 paramètres $F(j,k,t)$, chaque réel $F(j,k,t)$ représentant l'énergie de dix secondes ($2N'=10$) du signal d'énergie $E(k,t)$ dans la bande de fréquence j .

20 De manière à rendre $F(j,k,t)$ indépendant de l'amplitude du signal qui peut être plus ou moins fort, on rapporte ces valeurs à une valeur de référence, en l'occurrence la valeur maximale de $F_M(j,k,t)$ pour tous les k et j considérés. On obtient ainsi $K \times J$ paramètres
25 $F(j,k,t)/F_M(j,k,t)$.

On calcule en outre toutes les $2N'$ secondes, la phase de l'énergie $E(k,t)$ dans chacune des bandes de fréquences j : $P(j,k,t)$.

- 5 Pour cela, on calcule l'argument de la transformée de Fourier de $E(k,t)$ dans chacune des bandes de fréquence j :

$$P(j,k,t) = \text{Arg} \left[\sum_{n=-N'}^{n=N'} E(k,t+n/f_c) \cdot h'(n/f_c) \cdot e^{-4i\pi n/N'} \right]$$

- 10 Comme précédemment, on rapporte ces valeurs à une valeur de référence, en l'occurrence la valeur de $P(j,k,t)$ pour la seconde bande de fréquences ($j=1$) considérée car la référence temporelle de l'échantillon est inconnue : on ne connaît pas l'origine des temps.

- 15 Pour cela, on calcule les phases rapportées $\phi(j,k,t)$ par les formules suivantes :

$$\phi(1,k,t) = P(1,k,t)$$

$$\phi(j,k,t) = P(j,k,t) - P(1,k,t) \cdot \frac{f(k)}{f(1)}, \text{ pour } k > 1.$$

- 20 où les $f(k)$ sont les fréquences centrales des canaux k .
On obtient ainsi $K \times J$ paramètres correspondant aux valeurs de la phase rapportée $\phi(j,k,t)$.

- 25 On peut également considérer d'autres paramètres notamment, les valeurs moyennes de l'énergie $E(k,t)$ sur $2N'$ secondes et ce pour chaque bande de fréquences j :
 $E(j,k,t)$.

L'ensemble de ces paramètres normalisés définissent à intervalles réguliers, une empreinte que l'on peut considérer comme un vecteur $V(x(t))$. L'ensemble des paramètres normalisés par exemple
 5 $F(j,k,t)/F_M$ et $P(j,k,t)-P(j,1,t)$ définissent toutes les S' secondes, une empreinte que l'on peut considérer comme un vecteur $V(x(t))$ à $2 \times K \times J$ dimensions ($2 \times 127 \times 51$ soit environ 13000 dans notre exemple), une dimension par paramètre, chaque vecteur caractérisant un extrait
 10 de $2N'$ secondes du signal sonore $x(t)$, 10 secondes dans notre exemple.

On réitère cette caractérisation toutes les S' secondes, toutes les secondes par exemple ($S'=1$).

Comme représenté figure 3, un signal $x(t)$ durant T
 15 secondes est finalement caractérisé par L vecteurs V , L étant environ égal à T/S' .

Pour un signal sonore durant 10 mn soit 600s, on obtient 600 vecteurs c'est-à-dire $600 \times 2 \times J \times K$ paramètres.

Ces vecteurs sont mémorisés dans la zone de
 20 stockage 10 d'une base de données hébergée sur un serveur ou sur un disque compact. On a représenté figure 4 l'ensemble des vecteurs V d'un signal ou d'une œuvre A par V_A , de même V_B pour une œuvre B, etc.

Il est souhaitable de réduire le nombre de
 25 composantes de ces vecteurs c'est-à-dire le nombre de paramètres de façon à obtenir un vecteur ou une empreinte de taille plus réduite en vue de son stockage dans la base de données. De plus, lorsqu'il s'agira de comparer l'empreinte d'un signal sonore inconnu avec
 30 celles de la base de données, il sera souhaitable que

le nombre de paramètres à comparer soit réduit pour que cette recherche soit effectuée rapidement.

Or ces paramètres n'apportent pas tous la même quantité d'information, certaines pouvant être
 5 redondantes voire inutiles. C'est pourquoi on sélectionne les paramètres les plus pertinents parmi tous les paramètres, par un calcul d'information mutuelle présenté notamment dans la publication PROC. ICASSP'99, Phoenix, Arizona, USA, Mars 1999 H.YANG, S.VAN VUUREN, H.HERMANISKY, « Relevancy of Time-Frequency Features for Phonetic Classification Measured
 10 by Mutual Information ». On limite ainsi K à K_1 et J à J_1 .

On va présenter une méthode de sélection de ces
 15 paramètres.

Chacune des empreintes de ces signaux sonores c'est-à-dire chacun de ces vecteurs est classé dans un espace R à N dimensions, N étant le nombre de
 composantes des vecteurs. Pour des raisons de
 20 simplification, un exemple de classification pour des vecteurs à 2 dimensions P_1 et P_2 est représenté figure 5.

On définit des classes $C(m)$ regroupant les vecteurs par proximité, m variant de 1 à M . On peut par
 25 exemple décider qu'une classe correspond à une œuvre musicale : dans ce cas M est le nombre d'œuvres musicales mémorisées dans la base de données.

Il résulte du calcul de l'information mutuelle entre ces classes $C(m)$ et les paramètres, que la
 30 pertinence des paramètres est liée aux distances inter et intra classes : des paramètres pertinents

garantissent des distances inter-classes d relativement grandes comparées aux distances intra-classes D .

En ne retenant que les paramètres pertinents, on définit ainsi K_1 et J_1 .

5 On peut par exemple considérer cinq ($K_1=5$) bandes de fréquences respectivement centrées sur 344 Hz, 430 Hz, 516 Hz, 608 Hz et 689 Hz.

Des essais ont été effectués en prenant $J_1 = 3$.

10 Les classes $C(m)$ sont alors constituées à partir des vecteurs $V_q(x)$ ne comportant plus que $2 \times K_1 \times J_1$ composantes.

On va donner un exemple pour $K_1=5$ et $J_1=3$, de la taille mémoire d'une base de données contenant 1000 heures de musique et en considérant comme paramètres
15 $E(k,t)$ et $F(j,k,t)$, chacun de ces paramètres étant codé sur 4 octets.

Les paramètres $E(k,t)$ calculés toutes les 10 ms occupent $1000 \times 3600 \times 100 \times 5 \times 4$ octets soit environ 7 Giga octets.

20 Les paramètres $F(j,k,t)$ calculés toutes les secondes occupent $1000 \times 3600 \times 3 \times 5 \times 4$ octets soit environ 200 Méga octets.

Ces paramètres sont associés aux références des signaux sonores ; si l'on considère que les références
25 contiennent 100 caractères codés chacun sur un octet, ces références occupent $1000 \times 10 \times 100$ octets soit environ 1 Méga octet.

Une telle base de données occupera finalement environ 7 Giga octets.

Lorsqu'on souhaite identifier un signal sonore inconnu, on en établit d'abord l'empreinte, référencée $V(x_{inc})$ à la figure 6, comme décrit précédemment, sachant que le signal sonore inconnu peut être une œuvre musicale complète ou un extrait de cette œuvre.

La recherche de la classe de cette empreinte dans la base de données consiste alors selon une méthode classique illustrée figure 6, à comparer les paramètres de cette empreinte $V(x_{inc})$ à ceux des empreintes de la base de donnée. Les empreintes les plus proches, dites les plus proches voisins, définissent la classe de la façon suivante : la classe est celle de la majorité des plus proches voisins.

Un serveur de base de données 1 est schématiquement représenté figure 7. Il comprend une zone de stockage 10 des données de la base dans laquelle sont mémorisées les empreintes des signaux sonores assorties de leurs références. Il comprend en outre une mémoire 11 dans laquelle sont stockés les programmes de caractérisation et de recherche précédemment décrits, un processeur 12 assorti de mémoires de travail pour mettre en œuvre ces programmes. Il comprend bien sûr une interface d'entrée-sortie 13 et un bus 14 reliant ces divers éléments entre eux.

Lorsqu'il s'agit de rentrer dans la base de données 1, de nouveaux signaux sonores, l'interface 13 reçoit le signal $x(t)$ assorti de ses références ; s'il s'agit seulement d'un signal inconnu à identifier, l'interface 13 reçoit seulement le signal $x(t)$ inconnu.

En sortie, l'interface 13 fournit une réponse à la recherche d'un signal inconnu. Cette réponse est négative si le signal inconnu n'existe pas dans la zone de stockage 10 ; si le signal a été identifié, la
5 réponse comporte les références du signal identifié.

10

15

REVENDICATIONS

1. Procédé de caractérisation selon des paramètres spécifiques, d'un signal sonore $x(t)$ évoluant selon le temps t pendant une durée D dans différentes bandes de fréquences k et alors noté $x(k,t)$, caractérisé en ce
5 qu'il consiste à mémoriser le signal $x(t)$, à calculer l'énergie $E(k,t)$ dudit signal $x(k,t)$ pour chacune desdites bandes de fréquences k , k variant de 1 à K et selon une fenêtre temporelle $h(t)$ d'une durée $2N$, mémoriser les valeurs de l'énergie $E(k,t)$ obtenues, ces
10 valeurs constituant les paramètres spécifiques d'un extrait d'une durée $2N$ du signal sonore $x(t)$ et à réitérer ce calcul à intervalles réguliers pour obtenir l'ensemble des paramètres spécifiques pour la durée D du signal sonore $x(t)$.

15

2. Procédé selon la revendication précédente, caractérisé en ce qu'il consiste à calculer et mémoriser l'énergie $F(k,j,t)$ de $E(k,t)$ pour des bandes de fréquences j , j variant de 1 à J , selon une fenêtre
20 temporelle $h'(t)$ d'une durée $2N'$, les $J \times K$ valeurs de l'énergie $F(j,k,t)$ obtenues constituant les paramètres spécifiques d'un extrait d'une durée $2N'$ du signal sonore $x(t)$ et à réitérer ce calcul à intervalles réguliers pour obtenir l'ensemble des paramètres
25 spécifiques pour la durée D du signal sonore $x(t)$.

3. Procédé selon l'une des revendications précédentes, caractérisé en ce qu'il consiste à calculer la phase rapportée $\phi(j,k,t)$ de l'énergie $E(k,t)$ pour des bandes de fréquences j , j variant de 1 à J , et à inclure les valeurs de la phase rapportée $\phi(j,k,t)$ obtenues parmi les paramètres spécifiques du signal sonore $x(t)$.

4. Procédé selon l'une des revendications précédentes, caractérisé en ce qu'il consiste à calculer pour chaque bande de fréquence j , la valeur moyenne de l'énergie $E(k,t)$ sur $2N'$ secondes, à réitérer ce calcul à intervalles réguliers pour obtenir l'ensemble des paramètres spécifiques pour la durée D du signal sonore $x(t)$ et à inclure les valeurs moyennes obtenues parmi les paramètres spécifiques du signal sonore $x(t)$.

5. Procédé selon l'une des revendications précédentes, caractérisé en ce qu'il consiste à considérer les paramètres spécifiques d'un signal sonore $x(t)$ comme les composantes d'un vecteur représentatif de $x(t)$, à positionner les vecteurs dans un espace à autant de dimensions que de paramètres, à définir des classes regroupant les vecteurs les plus proches et à enregistrer lesdites classes.

6. Procédé selon la revendication précédente, caractérisé en ce que les classes présentent des distances inter-classes et des distances intra-classes et en ce qu'il consiste à sélectionner parmi les

paramètres spécifiques, les paramètres permettant d'obtenir des distances inter-classes relativement grandes devant les distances intra-classes et à enregistrer les paramètres sélectionnés.

5

7. Dispositif d'identification d'un signal sonore, caractérisé en ce qu'il comprend un serveur de base de données comprenant des moyens pour la mise en œuvre du procédé de caractérisation d'un signal sonore selon des paramètres spécifiques, selon l'une quelconque des revendications précédentes et des moyens de recherche dudit signal sonore dans la base de données.

10

8. Dispositif selon la revendication précédente prise en combinaison avec la revendication 5 ou 6, caractérisé en ce que les moyens de recherche comprennent des moyens de reconnaissance de la classe à laquelle ledit signal sonore appartient et des moyens de comparaison, par la méthode du plus proche voisin, des paramètres spécifiques du signal sonore inconnu avec les paramètres spécifiques de la base de données.

15

20

25

$x(k_i, t)$

1/5

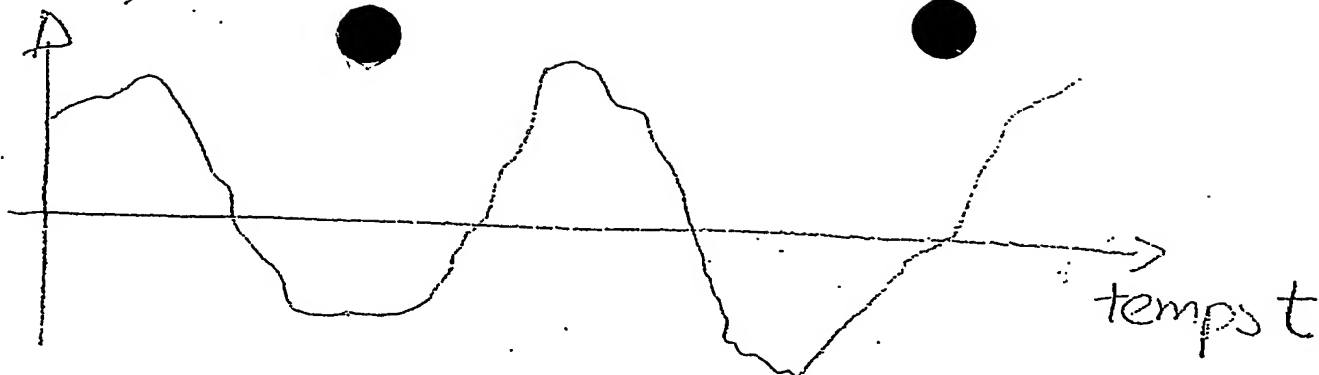


fig 1a)

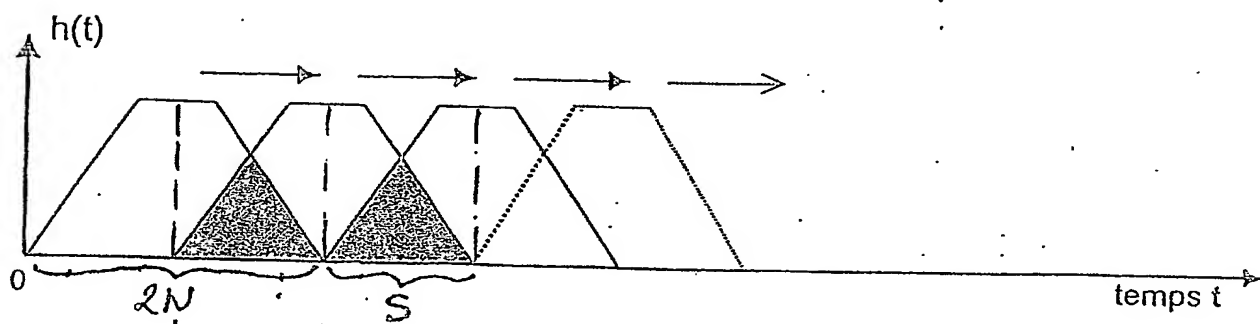


fig 1b)

$E(k_i, t)$

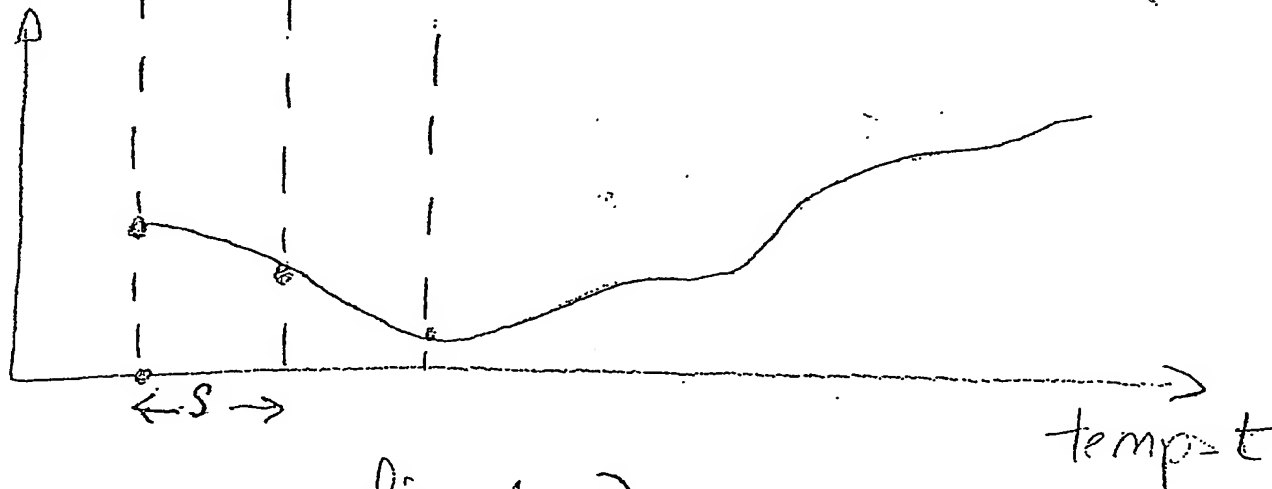


fig 1c)

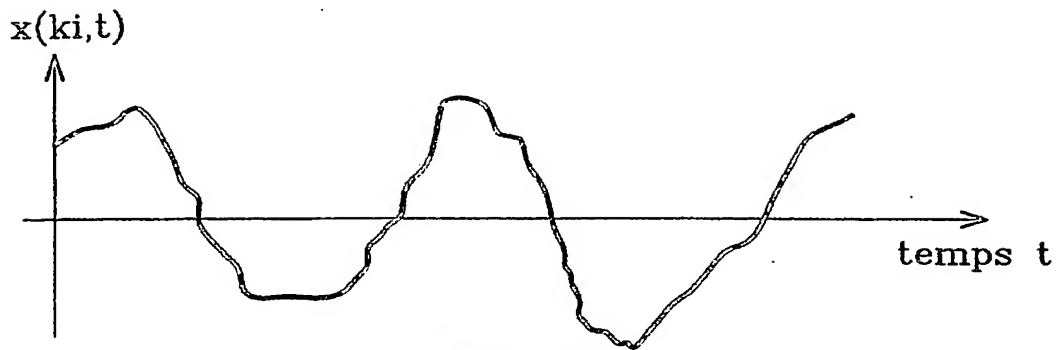


Fig. 1a

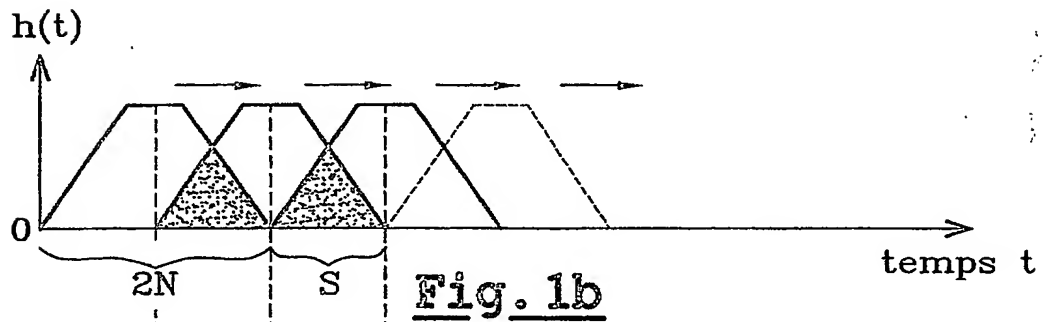


Fig. 1b

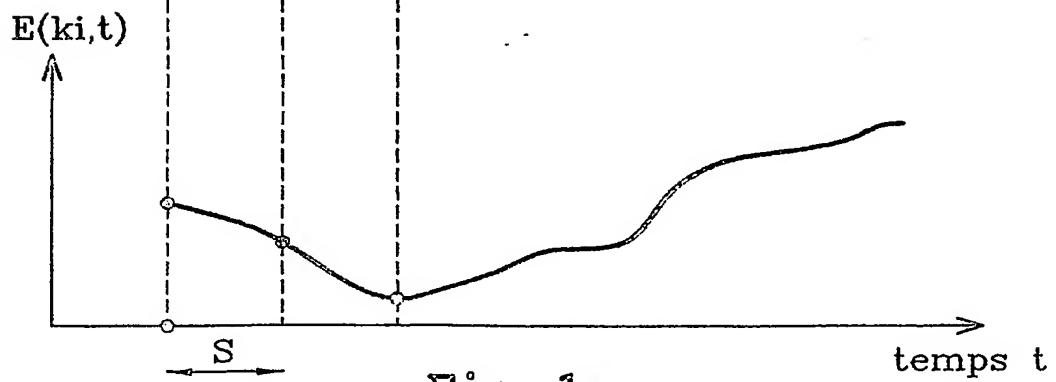


Fig. 1c

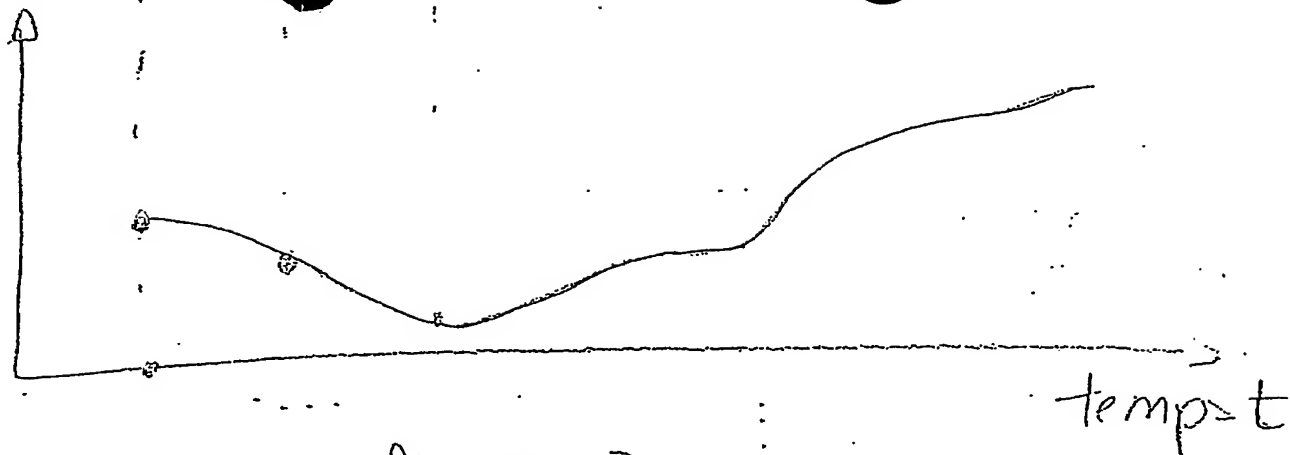
$E(k, t)$ 

fig 2 a)

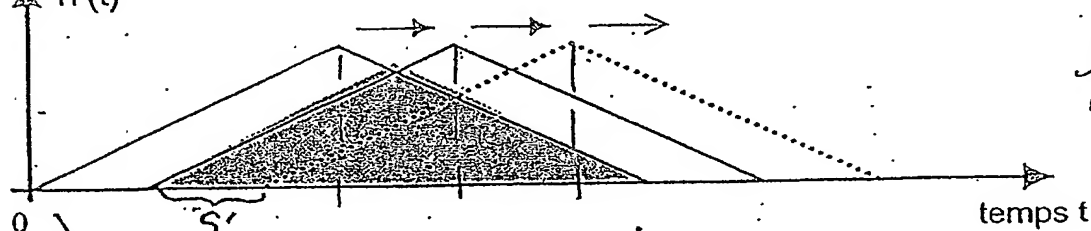
 $h'(t)$ 

fig 2b

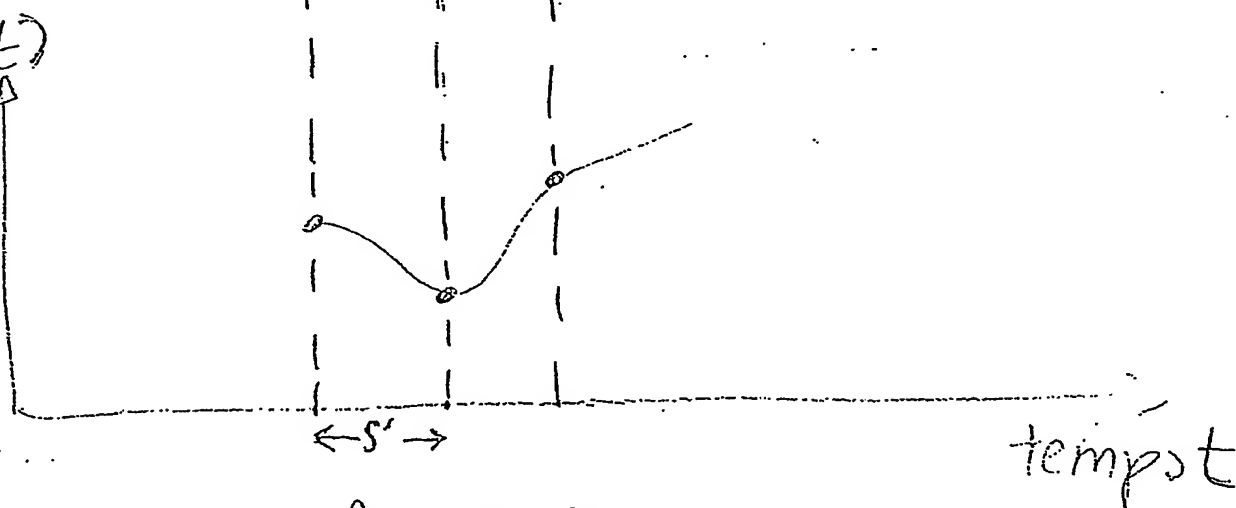
 $F(j_m, k_i, t)$ 

fig 2 c)

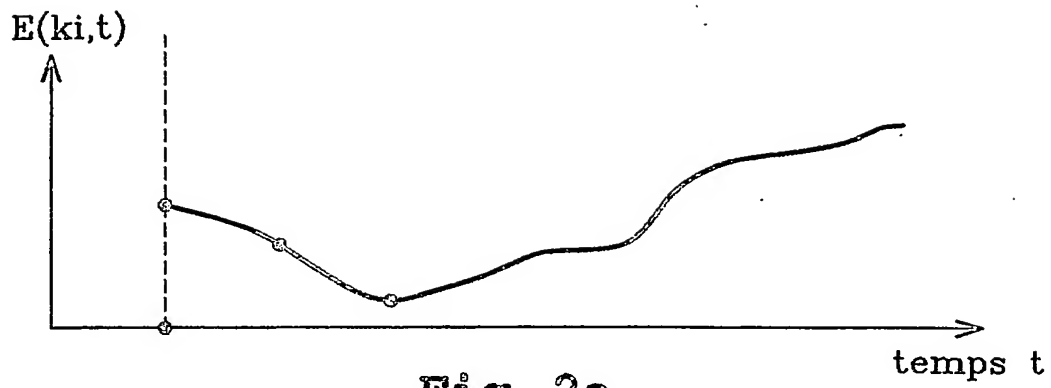


Fig. 2a

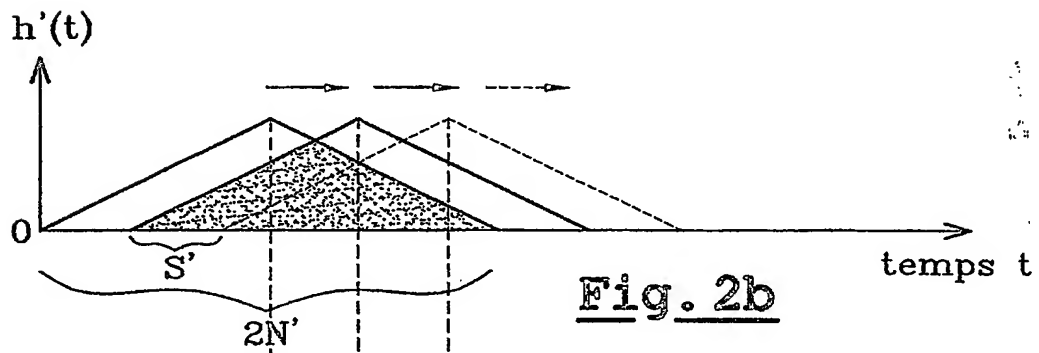


Fig. 2b

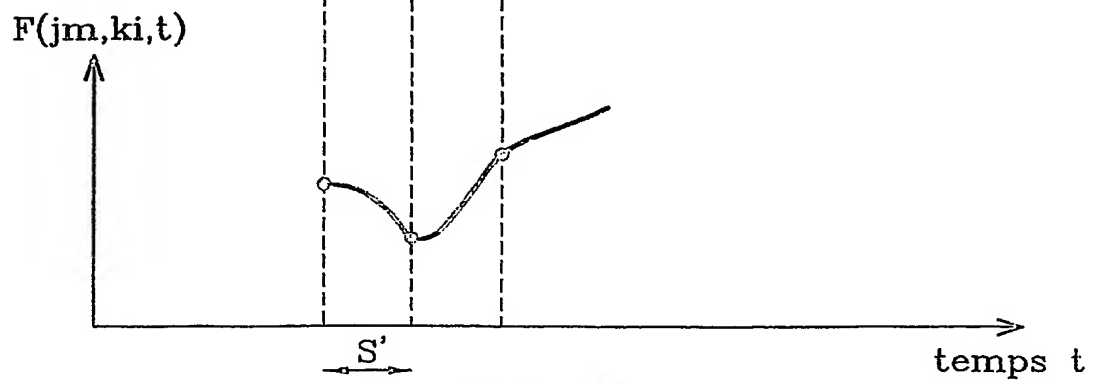


Fig. 2c

3/5

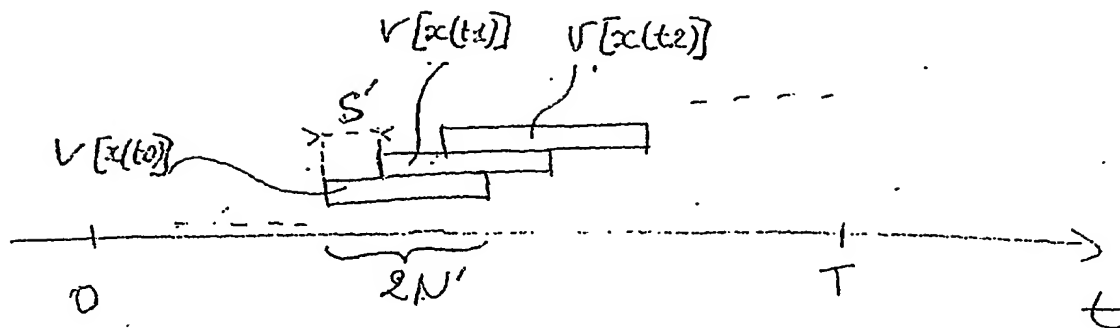


fig 3

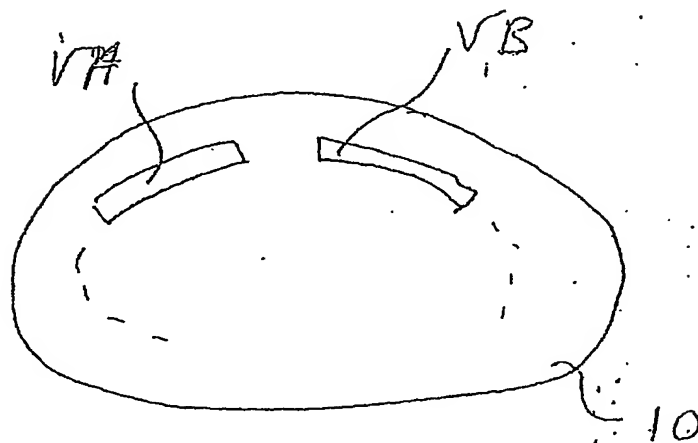


fig 4

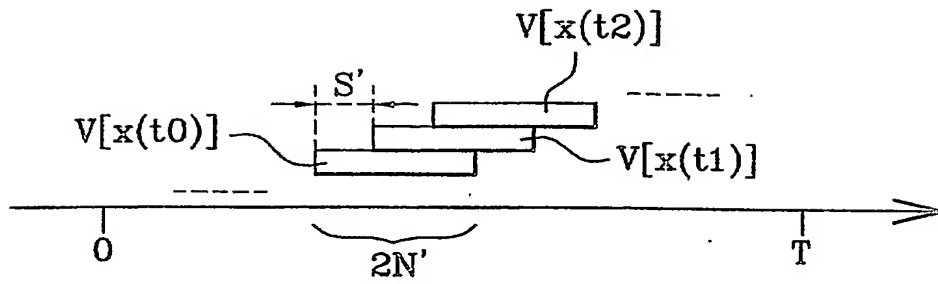


Fig. 3

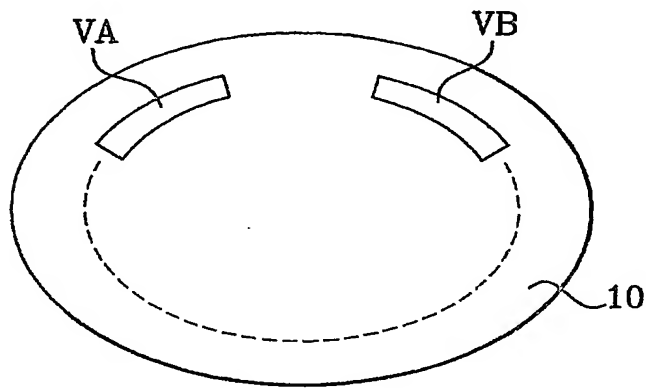


Fig. 4

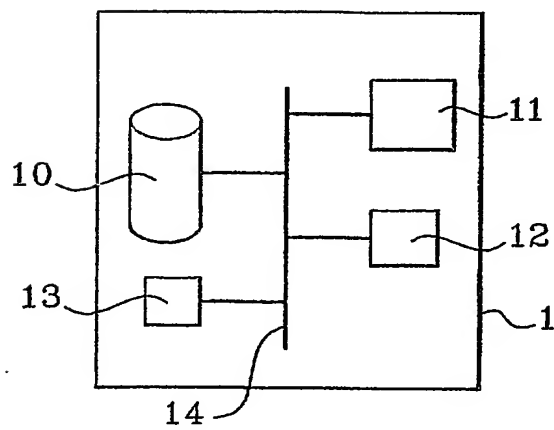


Fig. 7

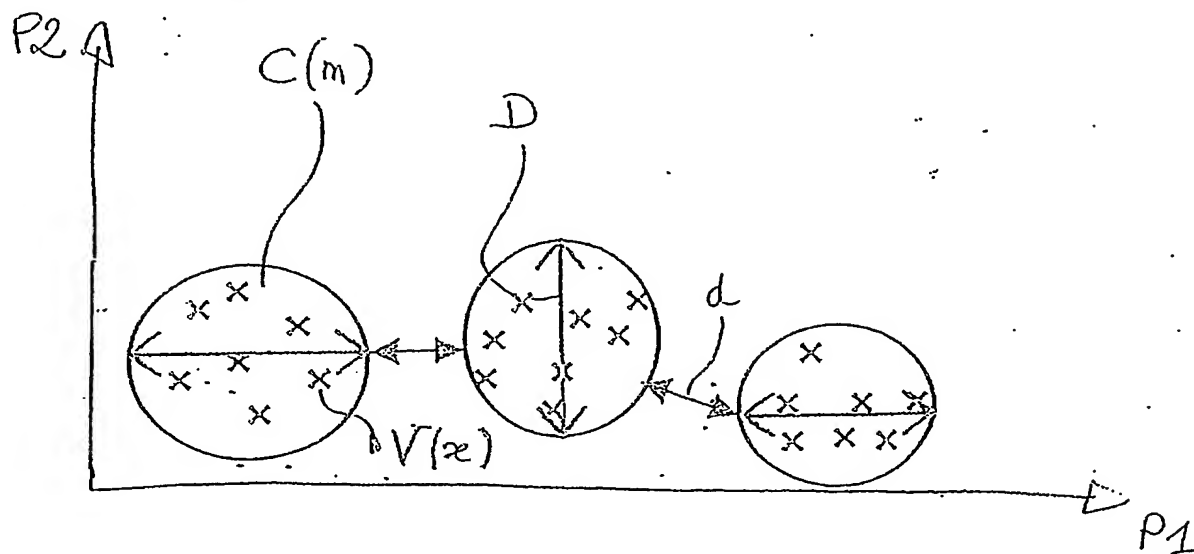


fig 5

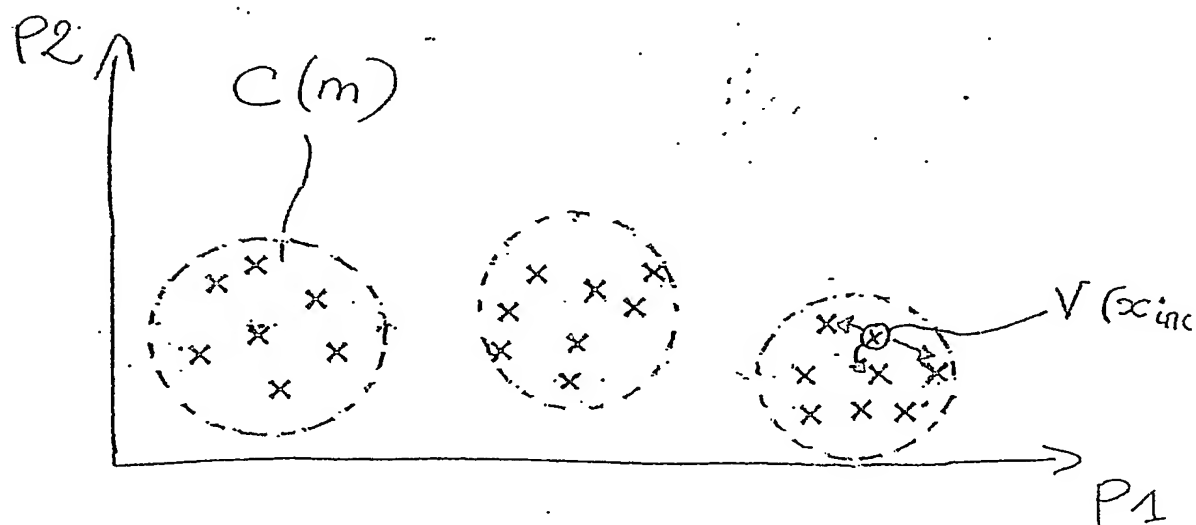
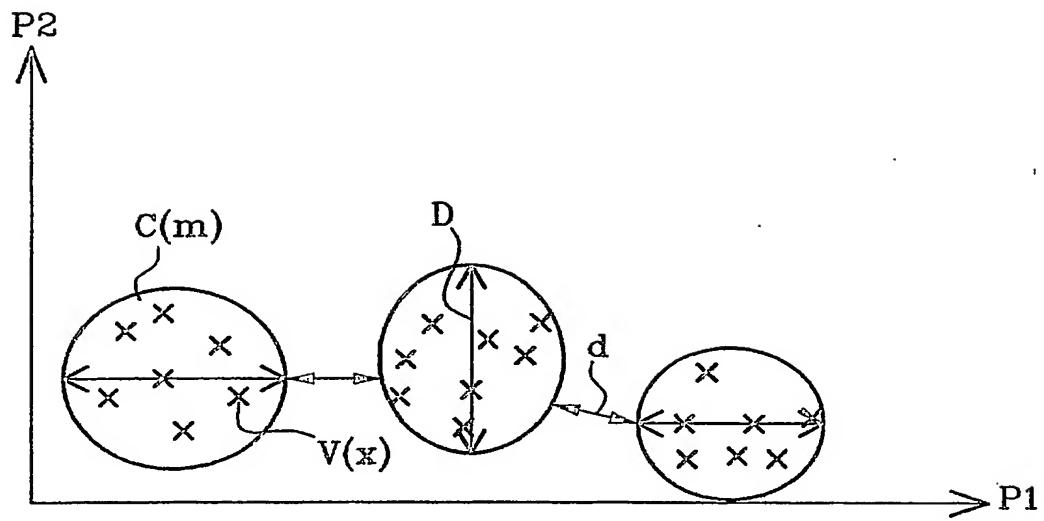
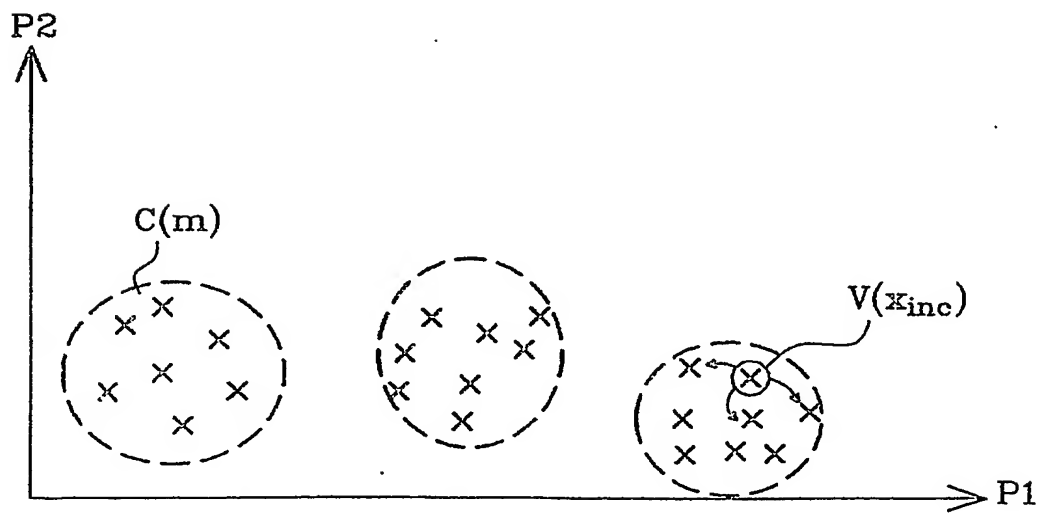


fig 6

Fig. 5Fig. 6

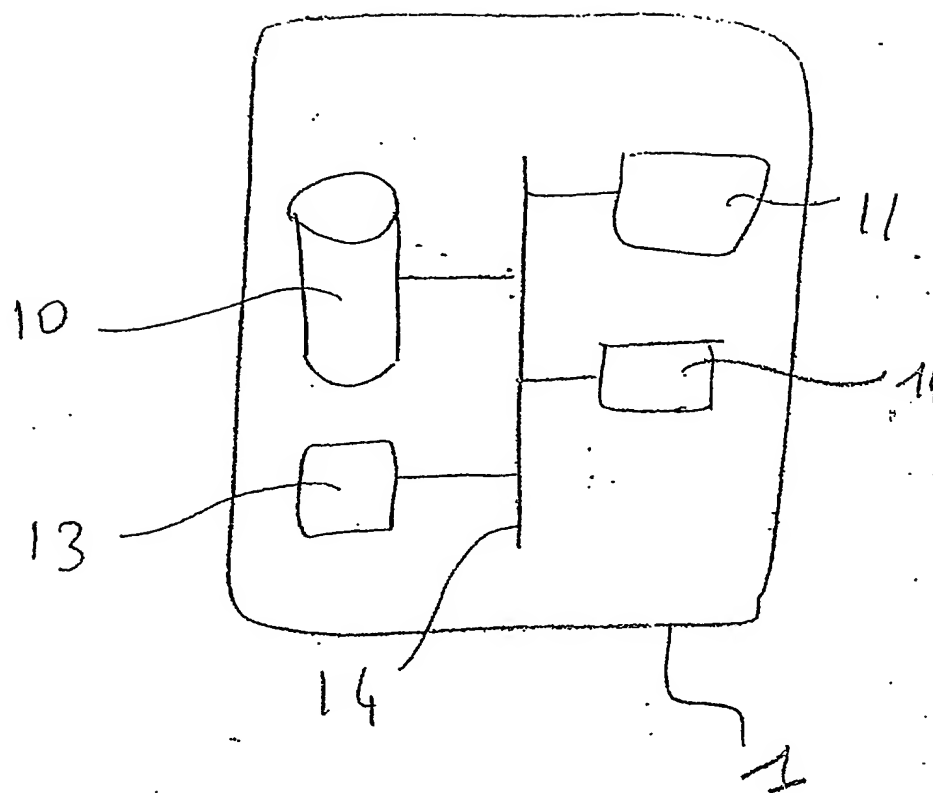


fig 7

DÉPARTEMENT DES BREVETS

26 bis, rue de Saint Pétersbourg

75800 Paris Cedex 08

Téléphone : 01 53 04 53 04 Télécopie : 01 42 93 59 30

DÉSIGNATION D'INVENTEUR(S) Page N° 1. / 1..

(Si le demandeur n'est pas l'inventeur ou l'unique inventeur)

Cet imprimé est à remplir lisiblement à l'encre noire

DB 113 W / 260899

Vos références pour ce dossier (facultatif)		015784 VG/CC	
N° D'ENREGISTREMENT NATIONAL		01 16949	
TITRE DE L'INVENTION (200 caractères ou espaces maximum) PROCÉDE DE CARACTERISATION D'UN SIGNAL SONORE			
LE(S) DEMANDEUR(S) : FRANCE TELECOM 6, Place d'Alleray 75015 Paris FRANCE			
DESIGNE(NT) EN TANT QU'INVENTEUR(S) : (Indiquez en haut à droite «Page N° 1/1» S'il y a plus de trois inventeurs, utilisez un formulaire identique et numérotez chaque page en indiquant le nombre total de pages).			
Nom		RODET	
Prénoms		Xavier	
Adresse	Rue	C/O Cabinet BALLOT 7, rue Le Sueur	
	Code postal et ville	75116	Paris
Société d'appartenance (facultatif)			
Nom		WORMS	
Prénoms		Laurent	
Adresse	Rue	C/O Cabinet BALLOT 7, rue Le Sueur	
	Code postal et ville	75116	Paris
Société d'appartenance (facultatif)			
Nom		PEETERS	
Prénoms		Geoffroy	
Adresse	Rue	C/O Cabinet BALLOT 7, rue Le Sueur	
	Code postal et ville	75116	Paris
Société d'appartenance (facultatif)			
DATE ET SIGNATURE(S) DU (DES) DEMANDEUR(S) OU DU MANDATAIRE (Nom et qualité du signataire) Paris, le 16 janvier 2002			
BALLOT Paul N° 92-1009			

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☒ **FADED TEXT OR DRAWING**
- ☒ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☒ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.